

Estimation of the Generalized Covariance Function
2. A Response Surface Approach

Nevada Univ., Las Vegas

Prepared for:

Environmental Monitoring Systems Lab., Las Vegas, NV

1987



U.S. Department of Commerce
National Technical Information Service

NTIS

TECHNICAL REPORT DATA <i>(Please read Instructions on the reverse before completing)</i>		
1. REPORT NO. 600/J-87/471	2.	3. RECIPIENT'S ACCESSION NO. PB89 208615/AS
4. TITLE AND SUBTITLE ESTIMATION OF THE GENERALIZED COVARIANCE FUNCTION: II. A RESPONSE SURFACE APPROACH		5. REPORT DATE
		6. PERFORMING ORGANIZATION CODE
7. AUTHOR(S) Thomas H. Starks, Allen R. Sparks		8. PERFORMING ORGANIZATION REPORT NO.
9. PERFORMING ORGANIZATION NAME AND ADDRESS Environmental Research Center, University of Nevada, Las Vegas, NV 89154; Computer Sciences Corporation, P.O. Box 15027, Las Vegas, NV 89114		10. PROGRAM ELEMENT NO.
		11. CONTRACT/GRANT NO.
12. SPONSORING AGENCY NAME AND ADDRESS Environmental Monitoring Systems Laboratory - LV, NV Office of Research and Development U.S. Environmental Protection Agency Las Vegas, NV 89193-3478		13. TYPE OF REPORT AND PERIOD COVERED
		14. SPONSORING AGENCY CODE
15. SUPPLEMENTARY NOTES		
16. ABSTRACT Delfiner (1976) suggested the use of a polynomial model for the generalized covariance function K and gave a regression procedure for estimation of coefficients of the polynomial. Starks and Fang (1982a) pointed out the lack of stability of these estimates because of the extremely skewed nature of the response variable in the regression and also because of the level of multicollinearity in predictor variables. (The problem of estimating the mean of a skewed distribution is also present in variogram estimation.) Kitanidis (1985) gave an unbiased quadratic estimation method for covariance functions that are linear in their parameters, which involves manipulation and inversion of large (n X n, where n is sample size) matrices. This linearity requirement eliminates the commonly used spherical and exponential models. Neuman and Jacobson (1984) used a stepwise iterative regression procedure for simultaneously estimating global drift and residual semivariogram. These authors state that their estimate of the residual semivariogram is a biased estimate. Cressie (1984) introduced a different approach to the spatial structure estimation problem. When data locations approximate a rectangular grid, he suggests the use of an additive two-factor analysis-of-variance model and median polish to remove drift and then the use of a robust variogram estimation procedure for the residuals.		
17. KEY WORDS AND DOCUMENT ANALYSIS		
a. DESCRIPTORS	b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group
18. DISTRIBUTION STATEMENT RELEASE TO PUBLIC	19. SECURITY CLASS (This Report) UNCLASSIFIED	21. NO. OF PAGES 18
	20. SECURITY CLASS (This page) UNCLASSIFIED	22. PRICE

EPA-600/J-87/471
PB89-208615

Estimation of the Generalized Covariance Function: II. A Response Surface Approach

Thomas H. Starks, Allen R. Sparks

Reprinted from Mathematical Geology, Vol. 19, No. 8, 1987

REPRODUCED BY
U.S. DEPARTMENT OF COMMERCE
NATIONAL TECHNICAL INFORMATION SERVICE
SPRINGFIELD, VA. 22161

Estimation of the Generalized Covariance Function: II. A Response Surface Approach¹

Thomas H. Starks² and Allen R. Sparks³

→ A procedure is proposed that employs first-moment estimation (kriging), cross-validation, and response surface analysis to estimate parameters of a generalized covariance function. Results from application of this procedure to two data sets are given. → to worksheet

KEY WORDS: Cross-validation, geostatistics, intrinsic random functions of order k , kriging.

INTRODUCTION

Delfiner (1976) suggested the use of a polynomial model for the generalized covariance function K and gave a regression procedure for estimation of coefficients of the polynomial. Starks and Fang (1982a) pointed out the lack of stability of these estimates because of the extremely skewed nature of the response variable in the regression and also because of the level of multicollinearity in predictor variables. (The problem of estimating the mean of a skewed distribution is also present in variogram estimation.) Kitanidis (1985) gave an unbiased quadratic estimation method for covariance functions that are linear in their parameters, which involves manipulation and inversion of large ($n \times n$, where n is sample size) matrices. This linearity requirement eliminates the commonly used spherical and exponential models. Neuman and Jacobson (1984) used a stepwise iterative regression procedure for simultaneously estimating global drift and residual semivariogram. These authors state that their estimate of the residual semivariogram is a biased estimate. Cressie (1984) introduced a different approach to the spatial structure estimation problem. When data locations approximate a rectangular grid, he suggests the use of an additive two-factor analysis-of-variance model and median polish to remove drift and then the use of a robust variogram estimation procedure for the residuals.

¹Manuscript received 9 April 1986; accepted 9 March 1987.

²Environmental Research Center, University of Nevada, Las Vegas, Nevada 89154.

³Computer Sciences Corporation, P.O. Box 15027, Las Vegas, Nevada 89114.

In most instances, estimation of the spatial structure (i.e., the generalized covariance function and the order k) of the intrinsic random function is a preliminary step to kriging. Kriging provides point (or block) estimates of the values assumed by the intrinsic random function in its current realization and, in addition, gives a measure of the precision of those estimates. Accuracy of this measure of precision depends on how well the estimated spatial structure approximates the true structure (at least locally). The typical approach in analysis of spatial structure has been to estimate directly second moments (i.e., semi-variograms or generalized covariance functions) of random variables. Estimation of second moments is less precise than estimation of first moments. Hence, imprecise second-moment estimation is performed to obtain the information needed to do first-moment estimation (kriging).

Cross-validation may be used to determine how well the estimated spatial structure locally approximates the true structure. Cross-validation here involves point kriging each data point x_α by using the estimated spatial structure and the data from neighboring points (but not from point x_α) to obtain an estimate $\hat{z}_\alpha$ of the observed value z_α and a kriging (estimation) variance $\sigma_{K\alpha}^2$ which, if the spatial structure model is correct, is the variance of the difference of the random variable Z_α associated with point x_α and the point kriging estimator $\hat{Z}_\alpha$. In cross-validation of a model, one looks at values of standardized residuals

$$r_\alpha = (\hat{z}_\alpha - z_\alpha) / \sigma_{K\alpha}$$

If the model is correct, residuals should behave as a sample from a population that has mean 0 and variance 1. Hence, if sample variance of the standardized residuals is radically different from 1, the spatial structure model is certainly suspect in that it is not properly measuring the precision of estimates of values at points where measurements are available and so cannot be expected to properly measure precision of estimates at points where measurements have not been taken. (Cross-validation was not employed to check the results of the proposed procedures in the last three papers mentioned in the first paragraph.)

Use of cross-validation in spatial structure analysis is not new; for example, cross-validation is used to compare models chosen in BLUEPACK's Automatic Structure Analysis subroutine (Delfiner et al., 1978), and another type of cross-validation is employed in that subroutine's "identification of the drift." Also, Davis and David (1978) suggested an iterative cross-validation procedure for estimation of generalized covariance functions. Several different models may give standardized residuals with sample means near 0 and sample variances near 1. Starks and Fang (1982b) made the conjecture that if the model is correct, the standardized residuals should behave as a sample from a population that has an approximately standard normal distribution. This conjecture was based on the fact that standardized residuals are linear contrasts of random variables. The conjecture may be used to discriminate between models satisfying the other cross-validation requirements.

The method introduced in this paper uses (first-moment) point kriging estimation, cross-validation, and response-surface analysis techniques to estimate spatial structure. This method in essence employs cross-validation of models to select a model.

PROCEDURE

An intrinsic random function Z of order k has been sampled at n locations x_α to yield a set of data points $\{(x_\alpha, z_\alpha)\}$. A general model for a generalized covariance function K for Z must be chosen. For purposes of illustration, the spherical model is considered. (However, the procedure applies in the same way to any appropriate model, and, indeed, several types of models may be fitted and compared for a set of data.) For the spherical model

$$K(h) = N\delta + S(1 - 1.5|h|/R + 0.5|h|^3/R^3) \quad \text{for } |h| \leq R \\ = 0 \quad \text{for } |h| > R$$

where N is the nugget effect, $(N + S)$ is the sill [i.e., $V(Z_\alpha)$], R is the range, and δ is a function of h that is 1 when $|h| = 0$, and 0 otherwise. To estimate the parameters, a search is made over a response surface (see Box and Wilson, 1951, or Box, Hunter, and Hunter, 1978, Chap. 15) to locate values of the parameters that give the "best fit" in point kriging the actual sample points using sets of nearest neighbors in the kriging. Kriging is performed by using $k = 2$ because $k = 2$ is usually sufficiently large (i.e., a complex drift function usually can be approximated locally by a quadratic drift) and also because an intrinsic random function of order 0 or 1 is also of order 2. Once parameters of the spherical model are estimated, kriging can be repeated with $k = 1$ and $k = 0$, and the smallest k can be used that does not degrade the fit.

In the response-surface analysis, the response is a measure of fit attained in kriging values at sample points by using observations from neighboring sample points, and the predictor variables, when estimating order of the drift and a spherical model, are the parameters k , N , R , and S . That is, for a given set of data, a response surface characterized by the equation

$$C = f(k, N, R, S)$$

is examined where C is a measure of goodness-of-fit of the spherical model and f is the function that characterizes that measure. A problem in this response-surface approach is the choice of a measure of fit for the kriging results. If one looks at the standardized residual

$$r_\alpha = (\hat{z}_\alpha - z_\alpha)/\sigma_{K\alpha}$$

the numerator is the difference between estimated and observed values at a point α and is thereby a generalized increment of order k . Thus the numerator is the

value of a random variable that has mean zero and standard deviation $\sigma_{K\alpha}$ if the model with the selected values of k , N , S , and R is (locally) correct. Because the numerator is a linear combination and is, in fact, a linear contrast of several values z_{α} , these standardized residuals reasonably may be expected to be distributed approximately as a sample from a standard normal distribution when the model is correct (Starks and Fang, 1982b, conjecture). Hence, a measure of deviation of standardized residuals from what is expected under the standard normal distribution may be employed as a measure of fit. Two such measures of fit are considered in this study, namely, mean squared deviation of the standardized residuals from their corresponding expected order statistics under the normal distribution, and mean absolute deviation of the standardized residuals from the expected order statistics. That is, if $r_{(\alpha)}$ represents the α smallest value of the n standardized residuals and $G_{(\alpha)}$ is the expected value of the α smallest observation in a sample of size n from a standard normal distribution, the proposed measures of fit (i.e., criteria for fit) are

$$CA = \sum |r_{(\alpha)} - G_{(\alpha)}|/n \quad \text{and}$$

$$CS = \sum \{r_{(\alpha)} - G_{(\alpha)}\}^2/n$$

Note that the right-hand side of each equation is a function of the values chosen for the model parameters through the obtained $r_{(\alpha)}$ values. The basic steps in the procedure are as follows:

1. Check data for outliers.
2. Check to determine whether variance is a function of the mean; if so, perform an appropriate transformation of the data to stabilize the variance.
3. Select a model for the generalized covariance function (e.g., spherical, white noise, exponential, etc.).
4. Specify the number of nearest neighbors that will be used in kriging each of the sample points in the selected subset.
5. Choose a starting point (i.e., a combination of values for the parameters of the generalized covariance function model); here $k = 2$ is used. Because often no evidence of spatial correlation is found, with real data after removal of drift, a small value for the range (but not as small as the shortest distance between points) is a wise starting point. An initial guess at the value of the sill might be a number somewhat smaller than the variance of the residuals (i.e., the residual mean square in the analysis of variance for the regression model) in a simple-model trend-surface analysis.
6. Use a grid search or the steepest-descent response-surface procedure to move to the combination of values for the parameters that gives a minimum value for the criterion employed.
7. Check the effect of reducing the value of k to 1 and to 0. Also consider other

general models for the generalized covariance function and compare minimum criterion (measure of goodness of fit) values found by the above steps for each general model.

CASE STUDY ONE

In this study, a data set previously discussed in Starks and Fang (1982b) that was obtained by adding a drift function

$$D(x, y) = 20[1 - \exp(-(x^2 + y^2)^{1/2}/4)]$$

to the value assigned to each of 225 points (x, y) by a turning bands simulation procedure is used. The points are on a 15 by 15 unit grid with coordinates going from (1, 1) to (15, 15). Experimental semivariograms for the turning bands data before and after the addition of the drift function are given (Fig. 1). The sample mean and variance of the turning bands data before addition of drift are 0.5111 and 1.9988; after adding drift, the corresponding statistics are 18.767 and 5.8627. Hence, after adding the drift function, the data represent observations on a nonstationary regionalized variable that has an approximate spherical correlation structure with a sill of about 2.2, a range somewhere between 5 and 6, and a near-zero nugget effect. The purpose of this study is to determine how

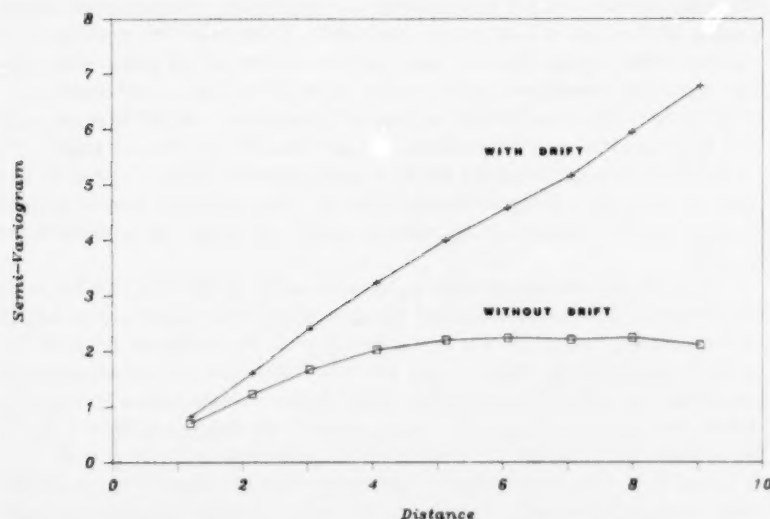


Fig. 1. Omni-directional experimental semivariograms for turning bands data with and without drift.

well this spatial correlation structure will be estimated when the response-surface procedure outlined in the previous section is applied to the with-drift data.

In this study, a simple grid search procedure is employed. At the start, k , N , and R are set at the values 2, 0, and 2, respectively. An arbitrary positive number is chosen for S ; in this case, $S = 5$. Each of the 225 sample points was kriged by using the nearest 12 neighbors to estimate a value $\hat{z}_\alpha$ for each point. The sample standard deviation s ($= 0.3696$) of the 225 standardized residuals r_α resulting from the kriging was calculated. Each of the standardized residuals was divided by s to obtain 225 numbers (having sample variance 1) that are the standardized residuals that would have been obtained if the same kriging had been done with an S value of $5s^2$ ($= 0.673$), where factor 5 is the initial S value. Criteria values $CA = 0.0648$ and $CS = 0.00652$ were calculated for this new set of standardized residuals. The range was increased by 1 to obtain $R = 3$, kriging was performed, and a new S value ($S = 1.180$) was determined in the same way so that standardized residuals for this model would have sample variance 1. Criteria values for these new standardized residuals were $CA = 0.0610$ and $CS = 0.00585$. Thus, reductions in values of both criteria were obtained by increasing the range by one unit and by taking appropriate values for the sills so as to obtain the desired property for the sample variance of the standardized residuals. A further increase of the range by 1 causes the range to be 4, and this presents a problem in that when kriging is based on the nearest 12 points (no two of which are more than four units apart), discrimination among models with ranges of 4 or more is impossible. To alleviate this problem, four subsets of the original data set were formed: namely, all 49 points with even first and second coordinates, all 64 points with odd first and second coordinates, all 56 points with even first and odd second coordinates, and all 56 points with odd first and even second coordinates. Again, the 225 points were kriged, but each point was kriged by using the 12 nearest neighbors from its subset. In this way, the maximum distance between pairs of points used in a kriging became 8 units, and now distinguishing between models with ranges up to about 6.5 is possible.

The process discussed above was run for values of R from 3 to 6.5 using increments of 0.1, and the data set was divided into four subsets as described in the previous paragraph. For $k = 2$ and $N = 0$, the minimum value of CA , 0.0467, was obtained with $R = 6.1$ and $S = 2.82$. Next, N was increased by increments of 0.05 with the process of looking at criterion values obtained for values of R between 3 and 6.5 repeated for each N value up to $N = 0.60$. In these trials, the minimum CA value (0.0437) was obtained for $N = 0.40$, $R = 5.2$, and $S = 1.26$. The nugget N cannot be estimated directly from the data with no assumptions being made. However, when an added assumption is made that the model has a specific form, such as the spherical, and when the distance between nearest neighbors is less than the range of influence, changes in N will

effect changes in the criteria values and can be used to improve the fit of the model to the data as has been done above. [When a value of N greater than 0 is employed, the procedure for selecting a value of S corresponding to a value of R must be changed because, for this case, the multiplication of a selected value of S by the sample variance of the standardized residuals will not produce a new S value for which standardized residuals will have unit sample variance. For $N > 0$, the first choice of S is multiplied by s^2 to obtain a new value for S as was done before, but now the kriging of the sample points must be repeated with the new value of S , and s must be recalculated for the new standardized residuals. If the new value of s is in the interval from 0.995 to 1.005, the new value of S is the one associated with R , and criteria values for CA and CS are calculated. If s is not within 0.005 of 1, the process is continued by multiplying the most recent values of S and s^2 to obtain a new S value which in turn is used in kriging to obtain new standardized residuals and a new s . Seldom is more than one iteration required. When the value of R is increased, a good starting value for S is obtained by multiplying the value determined for S , for the previous R value, and the ratio of the new value of R to the previous value of R .]

This process (using four subsets and values of R between 3 and 6.5) was repeated for $k = 1$. However, only nugget values of 0, 0.05, and 0.20 were employed because after running the first two, $N = 0$ and $N = 0.05$, the best results were evident for $N = 0$. The minimum CA value (0.0411) for $k = 1$ was obtained with $N = 0$, $R = 5.2$, and $S = 2.53$. Values of CA and CS change slowly near their minimums (Table 1) and the two criteria generally increase and decrease together. In addition to the minimum CA value, 0.04110 at $R =$

Table 1. Some Criteria Values Obtained for the Spherical Model

R	S	CA	CS
3.0	1.539	0.1425	0.02294
4.0	1.844	0.0603	0.00649
4.5	2.114	0.0486	0.00445
5.0	2.412	0.0415	0.00368
5.1	2.471	0.0413	0.00361
5.2	2.529	0.0411	0.00354
5.3	2.587	0.0412	0.00349
5.4	2.643	0.0412	0.00347
5.6	2.752	0.0412	0.00347
5.8	2.857	0.0414	0.00355
6.0	2.962	0.0421	0.00366
6.4	3.174	0.0448	0.00404

Note: $k = 1$, $N = 0$.

5.2, local minima for CA were obtained at $R = 5.5$ ($CA = 0.04120$) and at $R = 5.7$ ($CA = 0.04119$). For $k = 0$, the minimum CA value (0.0514) was obtained for $N = 0$, $R = 6.4$, and $S = 3.35$. This is a consequence of an almost linear variogram for the with-drift data (Fig. 1) and of the fact that only 2 of the 66 pairs of nearest-neighbor points are more than 6.4 units apart, whereas 10 are at least 6.3 units apart.

The results given above indicate that drift can be approximated locally (i.e., within the range of the 12 nearest neighbors) by a linear function (i.e., $k = 1$) of the coordinates; and that the nugget is 0, the range is somewhere between 5 and 6, and the sill is between 2.4 and 3, with our best specific estimate being $N = 0$, $R = 5.2$, and $S = 2.53$. The probability plot of the standardized residuals for this best estimate is given (Fig. 2). This same procedure was repeated on the original turning bands data (without drift), and the best CA result (0.0384) was obtained with $k = 0$, $N = 0$, $R = 5.3$, and $S = 2.4$, with

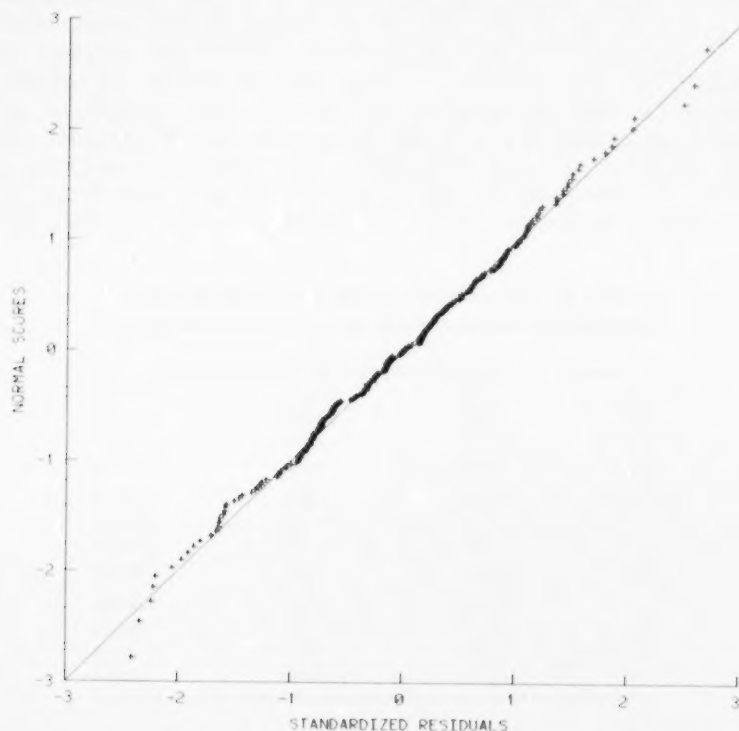


Fig. 2. Normal probability plot of 225 standardized residuals ($k = 1$, $N = 0$, $R = 5.2$, $S = 2.53$).

the same sort of slow change in criteria values near the minimum as observed (Table 1). (Actually the smallest CA value, 0.0381, for the turning bands data was attained for $k = 1$, $N = 0.05$, $R = 5.7$, and $S = 2.43$, but this was ignored because the criterion, when values were essentially the same for two values of k , was to choose the smaller value for k to obtain the simpler model.) Hence, for this example, at least, considerable internal consistency exists. In comparing results with the experimental semivariogram (Fig. 1), sill estimates appear large, but remember that the experimental semivariogram is also an estimate subject to error.

To determine how the number nn of nearest neighbors used in the kriging affects the fit of the model, two additional values, 8 and 20, were tried. For $nn = 20$, a minimum value, 0.04589, for CA was obtained for $k = 2$, $N = 0.3$, $R = 5.0$, and $S = 1.588$. The minimum criterion value is considerably larger in this case than when $nn = 12$, indicating that the spherical model with polynomial drift function has greater difficulty fitting the data over the larger region. This would seem especially true at the corners of the data set. For $nn = 8$, a minimum value for CA , 0.0539, was obtained for $k = 2$, $N = 0$, $R = 3.5$, and $S = 1.365$. This large increase in the minimum CA value may be due to the loss of information in removing four neighbors that were within the range of influence of the point being estimated and to the smaller number of pairs of neighbors with large intra-pair distances. The CS criterion gave the same estimates.

Would another model (e.g., an exponential model) possibly do better? One way of answering this question is to determine how small these criteria values might reasonably be expected to get. To do this, we generated 225 pseudo-random $N(0, 1)$ random numbers and calculated CA and CS . This was repeated 1000 times. Over the 1000 trials, the minimum values of CA and CS were 0.0302 and 0.0018, respectively. The corresponding fifth percentiles were 0.0435 and 0.0038, and the median values were 0.0710 and 0.0091. Hence, another model is unlikely to substantially reduce the criteria values.

CASE STUDY TWO

This study is based on coal ash data obtained from the Robena Mine Property in Green Co., Pennsylvania (Gomez and Hazen, 1970, Tables 19 and 20). Ash measurements (208) at core locations about 2500 ft apart on an approximately square grid with west coordinates greater than 64,000 ft were used. Drift is evident in a fairly steady increase in ash content from west to east over the 50 square mile region. One obvious outlier in the measured ash contents is 17.61%, whereas the other 207 measurements range between 7.00% and 13.07%. In the interest of obtaining stable estimates, this presumed outlier was removed from the data set before further analysis. The sample mean of the

remaining 207 measurements is 9.74% and the sample standard deviation is 1.16%. When ash measurements are regressed on a linear polynomial in the two location coordinates (i.e., simple trend surface analysis), the residual mean square is 1.008. This set of 208 ash measurements was used by Cressie (1984) to demonstrate several innovative exploratory data analysis techniques including robust estimation of the experimental semivariogram of residuals obtained through the use of median polish to remove drift.

To search for additional outliers that have undue influence on criteria values, all 207 points were kriged by using the white-noise model with $k = 2$, $N = 0.8$ and by employing the nearest 12 neighbors. The resulting value of s_r was 1.031. To obtain kriging results with s_r closer to 1, the data were kriged again with $N = 0.85 [= (0.8)(1.031)^2]$ (Table 2). The two points that gave the largest (in an absolute sense) standardized residuals (3.36 and -3.51) had ash concentrations of 8.69% and 13.06%. Location of the points are adjacent. We decided to treat them both as outliers and removed them. With the remaining 205 points, kriging gave smaller values for N and for the criteria (Table 2). The three points with the next largest standardized residuals (-2.84, -2.58, and -2.54) were removed, and kriging with the white-noise model again produced reductions in N and in the criteria. Removal of more points with large residuals did not reduce the values of the criteria.

On the basis of the heterogeneous nature of coal ash in coal, and from Cressie's (1984) results, the white-noise model was expected to give a good fit of the data, but the criteria values are considerably above the minimum values obtained in Case Study One for a similar number of data. For this reason, the nearest 8 neighbors were used instead of 12. This brought dramatic improvements in the criteria. Best results were obtained after removing only three outliers, 17.61%, 8.69% and 13.06%. [Results from four krigings with slightly

Table 2. Kriging Ash Data with the White-Noise Model

n	nn^a	N	Standard deviation	CA	CS
			(s_r)		
207	12	0.85	0.9998	0.08004	0.01489
205	12	0.73	1.0011	0.07078	0.00999
202	12	0.62	0.9966	0.05242	0.00509
207	8	1.02	0.9979	0.03940	0.00384
205	8	0.88	1.0012	0.03606	0.00228
205	8	0.89	0.9955	0.03505	0.00221
205	8	0.90	0.9900	0.03435	0.00220
205	8	0.91	0.9845	0.03446	0.00225

Note: nn = number of nearest neighbors used in kriging.

different values of N (Table 2) show how the criteria values change when s_r is close to 1.0.] Thus, although the quadratic ($k = 2$) interpolation based on 12 points did a good job for most points, it performed poorly for others; this would account for the small observed nugget effect but large criteria values.

To determine whether a spherical model would fit the data better than the white-noise model, the response-surface analysis using the set of 205 points and $nn = 8$ was performed. [To reduce the size of the coordinates, distance units were changed from feet to 10,000 ft, and this made the distance between grid points approximately 0.25 units.] The first factorial array consisted of four combinations of R and S values with R taking the values 0.30 and 0.45 and S taking the values 0.3 and 0.5. For each combination, the value of N was determined by trial and error (two trials on average) so as to obtain a value of s_r from the kriging that was within 0.005 of 1.000. This made results from different combinations of R and S comparable and reduced the dimension of the factor space by 1. Results for the four combinations (Table 3) show criteria values decreasing with decrease in R and with decrease in S . Because $R = 0.30$ is close to the typical distance between adjacent points in the array, little additional reduction is possible without having a de facto white-noise model. Hence, two more combinations ($R = 0.30$, $S = 0.10$) and ($R = 0.45$, $S = 0.10$) were tried with a smaller S value than was used in previous combinations. Results for these two combinations (Table 3) indicate that decrease in criteria values may be attained by reducing R and S . But at this stage, any reduction in R and S brings us to the white-noise model used previously. Hence, the response-surface analysis shows that the spherical model will not fit the data better than the previous white-noise model.

The final step is to check the model for $k = 1$. Kriging the 205 points by using the nearest eight neighbors, the white-noise model with $N = 0.82$ and $k = 1$ gives $s_r = 0.9974$, $CA = 0.07717$, and $CS = 0.01758$. Because these criteria values are far higher than those obtained with $k = 2$, the white-noise

Table 3. Kriging Ash Data with Spherical Models

N	R	S	Standard deviation	CA	CS
			(s_r)		
0.64	0.30	0.3	0.9954	0.03650	0.002389
0.46	0.30	0.5	1.0033	0.03843	0.002637
0.75	0.45	0.3	0.9962	0.03757	0.002524
0.65	0.45	0.5	1.0015	0.03914	0.002744
0.80	0.30	0.1	0.9992	0.03612	0.002301
0.84	0.45	0.1	0.9974	0.03632	0.002327

Note: $k = 2$, $n = 205$, $nn = 8$.

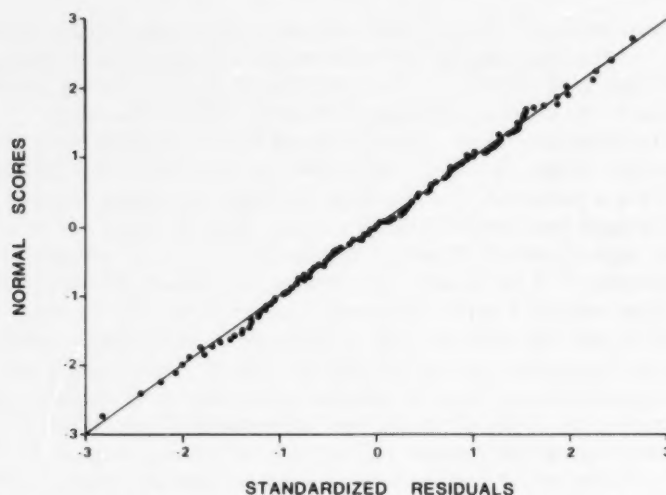


Fig. 3. Normal probability plot for 205 standardized residuals ($k = 2$, $nn = 8$, $N = 0.9$).

model with $N \approx 0.9$ and $k = 2$ was chosen as the model for the coal ash data. A normal probability plot of the standardized residuals obtained by kriging the 205 points with this model and $nn = 8$ is given (Fig. 3).

Cressie (1984) found a white-noise model with an N of approximately 0.65 (read from his robust variogram figure) for his residuals after median polish of the data. However, in a later paper, Cressie (1985) obtained a nugget value of about 0.75 when he made use of his robust procedure to calculate the east-west experimental semivariogram on residuals obtained from median polish.

COMMENTS AND CONCLUSIONS

The procedure given above for estimating parameters of a generalized covariance function is certainly not a method in which one can enter the data into the computer and then walk away a few seconds later with the answer. It requires interaction, and this in turn requires use of knowledge about the source of the data and of probable limits on the values of model parameters. As with most statistical procedures, the results get better as the amount of data available increases. Tracking both CA and CS is useful because, when they give inconsistent results, it usually indicates that something is wrong (e.g., outliers in the data).

When sample points are fairly uniformly distributed over the sampled region, the proposed estimation procedure provides information on the size of the neighborhoods over which the estimated model provides a reasonable approxi-

mation of the "true model." This is determined through the value of nn , number of nearest neighbors, for which small values of the criteria are obtained. The spatial structure model is recognized as unlikely to represent a region of interest in its entirety, but rather will give only reasonable approximations over small subregions (i.e., a polynomial drift function will approximate a complex drift only locally). This is one of the reasons why kriging is recommended to be done by using near neighbors of a point or block rather than the entire data set. However, the choice of the number of nearest neighbors to use in kriging has usually been by a rule of thumb rather than by being based on information on the size of the localities over which the model is a reasonable approximation. Kriging using the estimated spatial structure model should not employ a wider range of nearest neighbors than the range present in the sets of nn neighbors used in the spatial structure estimation.

Although this spatial structure estimation procedure is based on straightforward methods of point kriging, cross-validation, and response-surface analysis, it also is based on the unproven conjecture that under the correct spatial structure model, the standardized residuals r_α should have a distribution that is approximately standard normal. This conjecture is, in essence, a generalization of the Central Limit Theorem. Nevertheless, it seems reasonable and has worked well on the data to which it has been applied.

Because one cannot prove the hypothesis of the standardized residuals approximating a sample from a standard normal distribution when the correct model is used, one might compare models for which the standardized residuals have sample variance 1 by looking at their mean square errors (MSE), $\Sigma(z_\alpha - \hat{z}_\alpha)^2/n$. Although this is not an unreasonable approach, one objection to it is that it tends to give too much weight to the fitting of observations about which one has least supporting information (i.e., observations from sample points on the edge of the sampling or points having few close neighbors) and possibly to observations with large measurement errors where one tends to find the larger differences between predicted and observed values. Thus the model chosen by this minimum MSE procedure will tend to overfit the data about which one has the least collaborative information and underfit the data for which good collaborative information exists.

A referee has commented that the standardized residual $(z_\alpha - \hat{z}_\alpha)/\sigma_{K\alpha}$ is dangerous to use as a discriminator because it is the ratio of two random variables. Both numerator and denominator are functions of the data, in that model specifications are being changed from run to run of the procedure to bring the $\sigma_{K\alpha}$ values into line with the differences in the numerator. However, for a specified model, $\sigma_{K\alpha}$ is determined completely by sample locations and kriging neighborhood. Hence, when looking at the distribution of a set of standardized residuals, only the numerators are dependent on data values z_α . Therefore outliers in the z values do not have the same influence on standardized residuals here as they do in regression analysis where root mean square for error is af-

fects directly by outlier values. However, even in regression, common practice has been to look at the distribution of the standardized residuals to determine the appropriateness of the regression model.

In most environmental monitoring studies, quality assurance calls for duplicate pairs of samples and analysis of splits of samples so as to determine the contribution of procedural and analytical errors to the total variance. Measurement error variance (i.e., variance not due to spatial variation) should be kept separate from the nugget effect (short-range spatial variance) in calculating kriging standard deviations $\sigma_{K\alpha}$ (Dubrule, 1983). Unfortunately, this could not be done in Case Study Two because data required for estimation of measurement error variance were unavailable to the authors.

Results from the turning bands data in Case Study One may be expected to be somewhat better than results obtained from real data in that here one does not have to worry about the presence of outliers that are due to measurement errors, and, in addition, the turning bands procedure of averaging to produce the data gives the data an approximately multivariate normal distribution which thereby aids the normality assumption used in our criteria of fit. However, in fact, the minimum criteria values obtained for this case were not as small as those obtained in Case Study Two where real data were employed. We used the turning bands data for two reasons. First, it allowed us to add a drift function to stationary data and to determine whether the estimation procedure would provide a good estimate of the underlying spatial correlation, and, second, it allowed us to apply the method to spatially correlated data so as to observe whether the procedure would identify the spatial correlation or would treat it as a manifestation of drift and subsequently would identify the generalized covariance function to be a white-noise function. This second reason was a real worry in that in all of our applications of this procedure to real data sets, the estimated generalized covariance functions have been white-noise functions.

NOTICE

Although the research described in this article has been funded wholly or in part by the United States Environmental Protection Agency through Cooperative Agreement CR 812189-01 to the Environmental Research Center of the University of Nevada, Las Vegas, it has not been subject to Agency review and therefore does not necessarily reflect the views of the Agency, and no official endorsement should be inferred.

ACKNOWLEDGMENTS

The authors thank Mr. George Flatman and Dr. Evangelos Yfantis for their encouragement and helpful discussions of the project. The reviewers of this paper are also thanked for many helpful comments and criticisms.

REFERENCES

- Box, G. E. P., Hunter, W. G., and Hunter, J. S., 1978, *Statistics for experimenters*: John Wiley & Sons, New York, 653 p.
- Box, G. E. P. and Wilson, K. B., 1951, On the experimental attainment of optimum conditions: *J. Roy. Stat. Soc. Ser. B*, v. 13, p. 1-45.
- Cressie, N., 1984, Towards resistant geostatistics, in G. Verly (Ed.), *Geostatistics for natural resources estimation (Part 1)*: D. Reidel, Dordrecht, Holland, p. 21-44.
- Cressie, Noel, 1985, Fitting variogram models by weighted least squares: *Math. Geol.*, v. 17, p. 563-586.
- Davis, M. W., and David, M., 1978, Automatic kriging and contouring in the presence of trends: *J. Can. Pet. Tech.*, v. 17, no. 1, p. 1-10.
- Delfiner, P., 1976, Linear estimation of non-stationary spatial phenomena, in M. Guarascio et al. (Eds.), *Advanced Geostatistics in the Mining Industry*: D. Reidel, Dordrecht, Holland, p. 49-68.
- Delfiner, P., Renard, D., and Chiles, J. P., 1978, *BLUEPACK-3D Manual*: Centre de Geostatistique, Fontainebleau, France.
- Dubrule, Olivier, 1983, Two methods with different objectives: Splines and kriging: *Math. Geol.*, v. 15, p. 245-257.
- Gomez, M. and Hazen, K., 1970, Evaluating sulfur and ash distribution in coal seams by statistical response surface regression analysis: U.S. Bureau of Mines Report, RI 7377.
- Kitanidis, Peter K., 1985, Minimum-variance unbiased quadratic estimation of covariances of regionalized variables: *Math. Geol.*, v. 17, p. 195-208.
- Neuman, Shlomo P. and Jacobson, Elizabeth A., 1984, Analysis of nonintrinsic spatial variability by residual kriging with application to regional groundwater levels: *Math. Geol.*, v. 16, p. 499-521.
- Starks, T. H. and Fang, J. H., 1982a, On the estimation of the generalized covariance function: *Math. Geol.*, v. 14, p. 57-64.
- Starks, T. H. and Fang, J. H., 1982b, The effect of drift on the experimental semivariogram: *Math. Geol.*, v. 14, p. 309-320.